

# Theory of inter-edge superexchange in zigzag edge magnetism

J. Jung,<sup>1,\*</sup> T. Pereg-Barnea,<sup>2</sup> and A. H. MacDonald<sup>1</sup>

<sup>1</sup>*Department of Physics, University of Texas at Austin, USA*

<sup>2</sup>*Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA*

A graphene nanoribbon with zigzag edges has a gapped magnetic ground state with an antiferromagnetic inter-edge superexchange interaction. We present a theory based on asymptotic properties of the Dirac-model ribbon wavefunction which predicts  $W^{-2}$  and  $W^{-1}$  ribbon-width dependencies for the superexchange interaction strength and the charge gap respectively. We find that, unlike the case of conventional atomic scale superexchange, opposite spin-orientations on opposite edges of the ribbon are favored by both kinetic and interaction energies.

**Introduction**— Motivated by the seminal theoretical work of Kobayashi, Fujita, Wakabayashi and collaborators[1, 2, 3, 4], and by progress in graphene preparation[5], researchers have recently[6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17] reexamined the intriguing physics of edge magnetism in zig-zag terminated graphene nanoribbons from a number of different points of view. The magnetic state is a consequence of the nearly-flat subbands which occur at the Fermi level in a neutral zig-zag ribbon, and of the orbital character of the wavefunctions associated with these bands. Although there is still no conclusive[18] experimental evidence that a one-dimensional (1D) magnetic state occurs in ideal zigzag ribbons, the theoretically predicted state seems likely given that quite different electronic structure theories (from crude Hubbard models to elaborate *ab initio* DFT calculations) yield consistent[14] predictions and that there are at present no other ideas on how the unusual flat bands could be accommodated in the many-electron state. Present ribbons are far from ideal, however, and the main obstacle to realizing this paradigmatic example of  $d^0$  magnetism[19, 20] may lie in furthering recent progress[21, 22, 23] toward chemistry and defect control at the edge.

Mean-field-theory calculations predict that the ground state of a zigzag ribbon has unusually stiff parallel spin-alignment along each edge[13] and antiferromagnetic[15] inter-edge superexchange interactions. In this Letter we present a mostly analytic theory of the ribbon width  $W$  dependence of the important inter-edge superexchange interaction. Our theory relies on the properties of large  $W$  solutions of the continuum model approximation[7] for zigzag edges. We predict that interedge interactions can have a substantial influence on the properties of these unusual 1D magnets.

**Ribbon edge-state bands**— The  $\pi$ -orbital tight-binding model for a finite width ribbon yields a number of one-dimensional Bloch bands proportional to the ribbon width [16]. To a high degree of accuracy, zigzag magnetism is a rearrangement of only[14] the highest occupied ( $|k-\rangle$ ) and lowest unoccupied ( $|k+\rangle$ ) ribbon bands, whose transverse wavefunctions are respectively odd and even functions of carbon atom sites across the ribbon. Since the exchange physics which favors magnetism is local, it is revealed most clearly by forming states in this Hilbert space which are localized as far as possible at one

edge or the other:

$$|kL\rangle = \frac{1}{\sqrt{2}}(|k-\rangle + |k+\rangle), \quad |kR\rangle = \frac{1}{\sqrt{2}}(|k+\rangle - |k-\rangle) \quad (1)$$

where we have chosen the band transverse wavefunction amplitudes to be positive at the left most atom. These L(left), R(right) basis states can be expanded in terms of amplitudes on atoms in the ribbon unit cell:  $|kL\rangle = L_{kl}|k l\rangle_B$ ,  $|kR\rangle = R_{kl}|k l\rangle_B$  with sums over sites  $l$  within the unit cell implied. It is readily verified that  $L_{kl}$  and  $R_{k'l'}$  are strictly localized on opposite sublattices so that  $L_{kl}R_{k'l'} = 0$  [24]. In this representation, the  $\pi$ -band tight-binding model Hamiltonian  $H_{TB}(k) = t(k)\tau_x$  where  $\tau_x$  is a Pauli matrix and the left-right tunneling amplitude  $t(k) = -E_{TB}^-(k) = E_{TB}^+(k)$ . The bands in the Brillouin zone  $-\pi/a \leq k \leq \pi/a$  have periodicity  $2\pi/a$  and have inversion symmetry so we can restrict our attention to the interval  $0 \leq k \leq \pi/a$ . Zigzag edge magnetism follows from the following tight-binding model property [7, 16]. For  $2\pi/3 + q_e a \leq |k|$ , where  $q_e = 1/W = 2/\sqrt{3}aN$ , the states  $|kL\rangle$  and  $|kR\rangle$  are exponentially localized near their respective edges and the left-right hopping amplitude  $t(k)$  decreases rapidly with increasing  $W$ . (Here  $a = 2.46\text{\AA}$  is the lattice constant and  $N$  is the number of atom pairs in the ribbon unit cell.) Over this region of wavevector  $|kL\rangle$  and  $|kR\rangle$  are proper edge states.

**Hubbard model mean-field theory**— In order to explain our theory of the superexchange interaction we briefly summarize the mean-field theory of the magnetic state, which is particularly simple in the Hubbard model case. It is instructive to contrast two different collinear magnetic solutions of the mean-field equations, an antiferromagnetic (AF) one in which spins have opposite orientations on opposite edges and a ferromagnetic (F) one in which spins have the same orientation on opposite edges. The two LR basis spin-dependent mean-field Hamiltonians are given (up to a common constant) by:

$$H_\sigma^{AF} = \begin{pmatrix} -\sigma\Delta^{AF} & t \\ t & \sigma\Delta^{AF} \end{pmatrix}, \quad H_\sigma^F = \begin{pmatrix} -\sigma\Delta^F & t \\ t & -\sigma\Delta^F \end{pmatrix} \quad (2)$$

where  $\sigma = +/-$  for  $\uparrow/\downarrow$  spin, the  $k$ -dependence of  $t$  and the self-consistent exchange potentials is implicit, and

$$\Delta^{AF}(k) = U \sum_l L_{kl}^2 \langle m_l \rangle_{AF}, \quad \Delta^F(k) = U \sum_l L_{kl}^2 \langle m_l \rangle_F. \quad (3)$$

In Eq.( 3)  $m_l = (n_{l\uparrow} - n_{l\downarrow})/2$  is the site-dependent spin-density and  $n_{l\sigma}$  is the spin-dependent mean occupation number at site  $l$ . (Both solutions have  $n_l \equiv (n_{l\uparrow} + n_{l\downarrow}) \equiv 1$ , a convenient property of neutral ribbons which can be traced to particle-hole symmetry in the paramagnetic bands.) These self-consistent solutions are illustrated for a  $N = 20$  zigzag ribbon in Fig.( 1). Both the F (even  $m_l$ ) and AF (odd  $m_l$ ) state solutions are self-consistent. The mean-field equations are closed by evaluating  $m_l$  using

$$\langle m_l \rangle_{AF} = \frac{a}{2\pi} \int_0^{\pi/a} dk (L_{kl}^2 - R_{kl}^2) P(k) \quad (4)$$

$$\langle m_l \rangle_F = \frac{a}{2\pi} \int_{k_c}^{\pi/a} dk (R_{kl}^2 + L_{kl}^2) \quad (5)$$

where

$$P(k) \equiv \frac{\Delta^{AF}(k)}{((\Delta^{AF}(k))^2 + t(k)^2)^{1/2}} \quad (6)$$

is the degree of left-right edge polarization of the AF mean-field states. In the AF case, local spin-polarization follows from the opposite left-right polarizations of  $\uparrow$  and  $\downarrow$  states whereas in the F case the left-right polarization vanishes and spin-polarization follows from double occupation of  $\uparrow$  bands for  $k_c < |k| < \pi/a$  where  $k_c$  is the wavevector at which the  $\uparrow, +$  band and the  $\downarrow, -$  band cross as illustrated in Fig. (1).  $\Delta^{AF}$  and  $\Delta^F$  are nearly (but not quite!) identical because their  $m_l$ 's differ mostly in the middle of the ribbon which has little influence on the edge states.

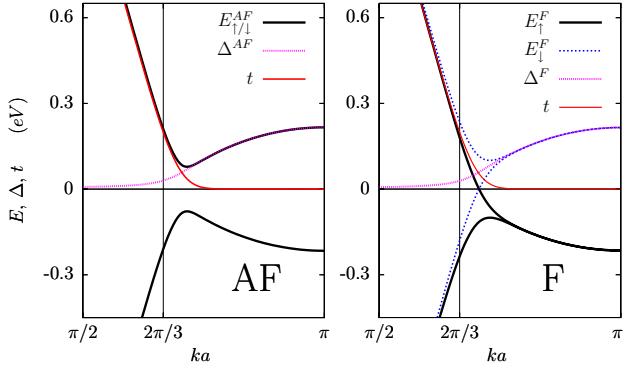


FIG. 1: (Color online) Hubbard model mean-field calculations for  $\gamma_0 = 2.6\text{eV}$  and  $U = 2\text{eV}$  for a zigzag ribbon with  $N = 20$ . *Left panel:* Mean-field energy bands for the AF state:  $E_{\sigma}^{AF}(k) = \pm (\Delta^{AF}(k)^2 + t^2(k))^{1/2}$  for both spins with the low-energy states of one spin concentrated on one side of the ribbon and the low-energy states of the opposite spin concentrated on the opposite side.  $t(k)$ , the left-right hopping parameter, is quite insignificant for this ribbon width in the edge state region.  $\Delta^{AF}(k)$  is dominant in the edge state region because of large local spin-polarizations. *Right panel:* Mean-field energy bands for the F state:  $E_{\sigma}^F(k) = \sigma \Delta^F(k) \pm |t(k)|$ . Note that  $\Delta^{AF}$  and  $\Delta^F$  are nearly identical. The bands are periodic with periodicity  $2\pi/a$  and inversion symmetric.

*Ribbon width scaling rules*— From solutions of the graphene continuum model [7, 16] we obtain for the region near  $k =$

$\pm(2\pi/3 + q)$  and small  $q$  the expression

$$t(\pm(q + 2\pi/3)) = (\sqrt{3}\gamma_0 a/2) \sqrt{q^2 - z^2}, \quad (7)$$

where  $z$  satisfies

$$\begin{aligned} zW \coth(zW) &= qW, & q > q_e \\ zW / \tan(Wz) &= qW, & q < q_e. \end{aligned} \quad (8)$$

In the continuum model

$$R_{kl}^2 \rightarrow R_k^2(y) = 2z(\cosh(2zy) - 1) / (\sinh(2Wz) - 2Wz) \quad (9)$$

for  $k_c < k < \pi/a$  and

$$R_k^2(y) = 4z \sin(zy) / (2Wz - \sin(2Wz)). \quad (10)$$

for  $0 < k < k_c$ . The left centered functions can be obtained through the symmetry relation  $L_{kl} = R_{k2N-l}$ . It follows that

$$t(k) = \frac{\gamma_0}{W} \tilde{t}(qW) \quad (11)$$

and that

$$R_k(y)^2 = W^{-1} \chi^2(qW, y/W), \quad (12)$$

where the functions  $\chi^2(x)$  and  $\tilde{t}(x)$  are implicitly defined by the above equations. We have verified that these scaling relations apply accurately even in quite narrow ribbons.

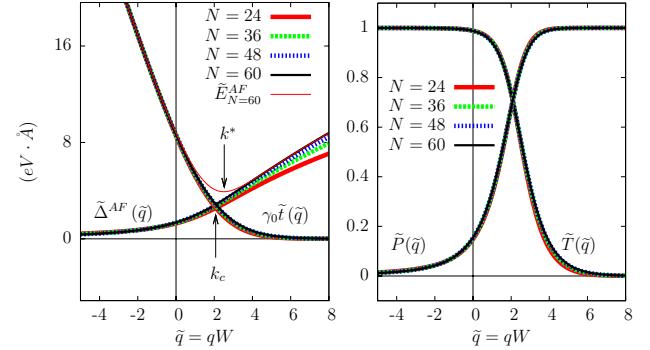


FIG. 2: (Color online) *Left panel:* Dependence of  $\tilde{\Delta}^{AF}$  and  $\tilde{\chi}_0 \tilde{t}$  on the scaled coordinate  $\tilde{q} = qW$  and the corresponding  $N = 60$  AF state quasiparticle bands. Note that the self-consistently calculated  $\tilde{\Delta}^{AF}$  approaches a well defined function at large  $N$ . The positions  $k_c$  and  $k^*$  are respectively the values of  $k$  at which  $\Delta^{AF} = t$  and the band gap minimum occurs. *Right Panel:* Scaling collapse of antiferromagnetic state self-consistent left-right polarization  $\tilde{P}$  and symmetric-antisymmetric polarization  $\tilde{T}$  represented in the scaled coordinate  $\tilde{q}$ . Note that  $P^2 + T^2 = 1$  by definition.

From Eq.(3) we see that the Hubbard model exchange potentials depend on local spin polarizations  $\langle m_l \rangle$  which are large only close to the edge and approach a well defined limit already for quite narrow ribbons; the form of the spin-polarization near each edge is a single-edge property unrelated to interedge interactions. From this observation and the above

scaling relations for the zigzag edge states, we propose the following scaling rule for the form of the exchange potential

$$\Delta^{AF/F}(k) = W^{-1} \tilde{\Delta}^{AF/F}(qW). \quad (13)$$

Since both  $\tilde{\Delta}(k)$  and  $\tilde{t}(k)$  depend only on  $qW$  it follows from Eq.( 6) that  $P(k)$  also depends only on  $qW$ . We have verified numerically that this scaling relationship holds accurately for sufficiently wide ribbons as illustrated in Fig.( 2).

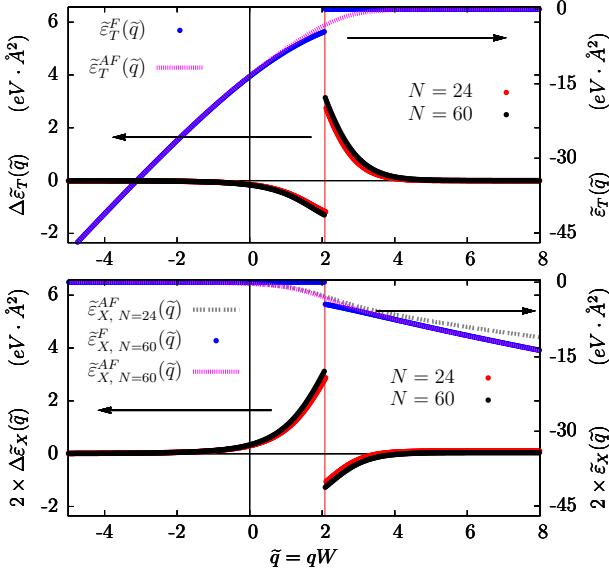


FIG. 3: (Color online)  $k$ -resolved contributions to the kinetic ( $\tilde{\varepsilon}_T$ ) (upper panel) and exchange ( $\tilde{\varepsilon}_X$ ) of the F and AF states and the corresponding  $F - AF$  differences as a function of the scaled momentum coordinate  $\tilde{q} = qW$ .  $\varepsilon_T^F$  and  $\varepsilon_T^{AF}$  are the integrands in the kinetic energy expression Eq.( 15), and  $\varepsilon_X^{F/AF}$  is the corresponding quantity for the exchange interaction energy. The discontinuities in the ferro case are due to the crossing between the majority-spin symmetric and minority-spin antisymmetric bands at  $k_c$ , indicated by a thin vertical line. Although the kinetic energies roughly double the interaction energies at most  $k$  values, the exchange contribution to superexchange is much larger because of weaker cancellation between  $|k| < k_c$  and  $|k| > k_c$  regions.

*Inter-edge interaction*— The strength of the superexchange interaction which determines the alignment between magnetization directions on opposite edges is given by the total energy difference between AF and F solutions. Because the electrostatic Hartree energies of both states are identical, the energy difference per edge carbon atom  $\Delta E$  can be separated into band (kinetic) and exchange energy contributions:

$$\Delta E = E^F - E^{AF} = \Delta T + \Delta E_X. \quad (14)$$

The difference of kinetic energies between AF and F solutions are determined by contributions from occupied edge band states:

$$\begin{aligned} T^{AF} &= -\frac{2a}{\pi} \int_0^{\pi/a} dk t(k) T(k) \\ T^F &= -\frac{2a}{\pi} \int_0^{k_c} dk t(k) \end{aligned} \quad (15)$$

where  $T(k) = (1 - P^2(k))^{1/2}$  is the symmetric-antisymmetric polarization of AF states. The F as well as non-interacting band eigenstates have  $T(k) \equiv 1$ . In the non-interacting ground state the lower energy state is fully occupied and the total energy contains all the band energy. Both AF and F states sacrifice band energy contributions in the region  $|k| \gtrsim 2\pi/3a$  in order to gain interaction energy. In the ferromagnetic case the band energy gain is sacrificed completely for  $|k| > k_c$ , the wavevector at which  $\Delta = t$ . At larger values of  $|k|$  both bonding and antibonding states are occupied for one spin and both are empty for the other spin. There is therefore an abrupt separation at  $|k| = k_c$  between wavevectors which contribute to band energy and regions which contribute to the exchange energy, discussed below. In the AF case, on the other hand,  $T$  crosses smoothly as a function of scaled wavevector  $\tilde{q} = qW$  from the kinetic energy contributing regime at small  $|k|$  to the exchange energy contributing regime at large  $|k|$ . The interedge interaction is due to this difference. Using the scaling properties of  $t$  and  $\Delta$  the kinetic energy contribution to the difference can be written as an integral over  $\tilde{q}$ :

$$\Delta T = \frac{2a\gamma_0}{\pi W^2} \left[ \int_{-\infty}^{\tilde{q}_c} d\tilde{q} \tilde{t}(\tilde{q}) [\tilde{T}(\tilde{q}) - 1] + \int_{\tilde{q}_c}^{\infty} d\tilde{q} \tilde{t}(\tilde{q}) \tilde{T}(\tilde{q}) \right]. \quad (16)$$

The integrals in Eq.( 16) converge at  $-\infty$  because  $\tilde{T}(\tilde{q})$  approaches 1 rapidly and at  $\infty$  because  $\tilde{T}(\tilde{q})$  approaches 0 rapidly. The contribution from  $\tilde{q} < \tilde{q}_c$  is negative while the contribution from  $\tilde{q} > \tilde{q}_c$  is positive. Substantial cancellation leads to a small kinetic energy contribution to  $\Delta E$ .

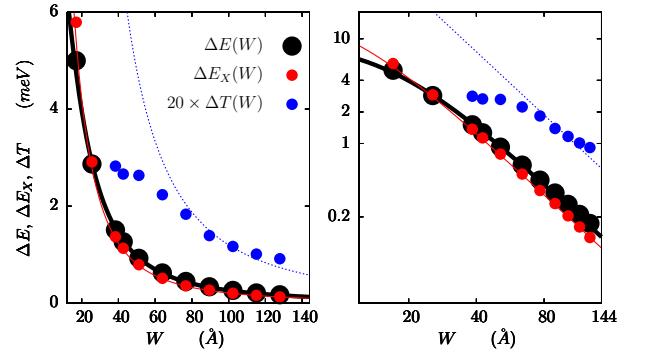


FIG. 4: (Color online) The difference between the F and AF states in total energy ( $\Delta E$ ), exchange energy ( $\Delta E_X$ ) and kinetic energy ( $\Delta T$ ) plotted on a linear scale (left) and a logarithmic scale (right). The total energy difference follows a  $W^{-2}$  decay law at large  $W$  and is dominated by exchange energy contribution. The kinetic energy contribution is substantially smaller and the asymptotic decay law develops only for sufficiently large ribbon width. The total energy  $\Delta E$  was fitted with  $2.7/(W^2 + 280)$ , the exchange energy  $\Delta E_X$  with  $2.1/(W^2 + 100)$  and the fitting for the kinetic contribution was obtained from the difference between both resulting in  $0.6/W^2$  (represented with  $\times 20$  magnification), all terms given in  $eV\text{\AA}^{-2}$  units. We used a 12K  $k$ -point sampling density in the 1D Brillouin zone.

The exchange energy integrands satisfy the scaling relations similar to the kinetic terms and we can write the exchange energy difference as

$$\Delta E_X = \frac{a}{\pi W^2} \left[ \int_{-\infty}^{\tilde{q}_c} d\tilde{q} \tilde{\Delta}^{AF}(\tilde{q}) \tilde{P}(\tilde{q}) + \int_{\tilde{q}_c}^{\infty} d\tilde{q} \tilde{\Delta}^{AF}(\tilde{q}) [\tilde{P}(\tilde{q}) - 1] \right] + \frac{1}{3} (\Delta^{AF}(\pi/a) - \Delta^F(\pi/a)) \quad (17)$$

The first two terms are similar to the corresponding band energy contributions, with the discontinuity at  $q_c$  again due to the band crossing in the ferromagnetic state. We write this contribution to the superexchange interaction as  $J_X/W^2$ . An additional contribution appears because  $\Delta^{AF}$  and  $\Delta^F$  are not quite identical for  $\tilde{q} \rightarrow \infty$ .

In the Hubbard model we can relate the asymptotic difference in  $\Delta$  to the difference in spin polarization on the edge atom:  $\delta\Delta \equiv \Delta^{AF}(\pi/a) - \Delta^F(\pi/a) = U (\langle m_{edge} \rangle_{AF} - \langle m_{edge} \rangle_F)$ . Labeling the leftmost site as site 1, noting that  $R_{k1}^2 = 0$  and recalling the definitions of  $\langle m_l \rangle$  in Eqs. (4,5) we find that

$$\begin{aligned} \delta\Delta &= \frac{a}{2\pi} \int_0^{k_c} dk L_{k1}^2 P(k) + \frac{a}{2\pi} \int_{k_c}^{\pi/a} dk L_{k1}^2 (1 - P(k)) \\ &\equiv \frac{3J_{\delta m}}{W^2} \end{aligned} \quad (18)$$

Adding the three contributions, the total superexchange interaction is

$$\Delta E(W) = W^{-2} (J_K + J_X + J_{\delta m}). \quad (19)$$

For  $\gamma_0 = 2.6\text{eV}$  and on site repulsion  $U = 2.0\text{eV}$  that results in band gaps similar to LDA [14], we find that the kinetic and exchange contributions to the interaction are  $J_K = 0.6\text{eV}\cdot\text{\AA}^2$ , and  $J_X + J_{\delta m} = 2.1\text{eV}\cdot\text{\AA}^2$ . (See Fig.( 4).) Separately  $J_{\delta m} \simeq 1.15\text{eV}\cdot\text{\AA}^2$  implying that the interaction contribution is composed in approximately equal measures of contributions from  $q$  near  $q_c$  and contributions far in the edge regime.

*Conclusions and discussions*— Our analysis shows that the antiferromagnetic inter-edge superexchange interaction in magnetic zigzag nanoribbons is the sum of three contributions (band-energy, exchange-energy, and edge spin-polarization) all of which arise from a region of the ribbons's 1D Brillouin-zone which is centered on  $|k| = 2\pi/3a$  and scales in width as  $1/W$ . Unlike the familiar case of atomic-scale superexchange interactions, in which antiferromagnetic spin-arrangements lower the kinetic energy at a cost in interaction energy, all three contributions have the same sign - with the kinetic contribution being substantially smaller in magnitude. Our conclusions rest primarily on analytic properties of continuum model solutions to the  $\pi$ -band model for zig-zag nanoribbons and depend on the particle-hole symmetry of graphene's conduction and valence bands. We have demonstrated numerically that the continuum model predictions are accurate, even in narrow nanoribbons. Although some details of our analysis depend on the simplified Hubbard model we use, we expect that the scaling properties of states near  $|k| = 2\pi/3a$  to be general and that our qualitative conclusions will apply to any mean-field-theory treatment of zigzag ribbon magnetism.

Collective spin-behavior is expected[13] to be important in zigzag ribbon magnets, even though they are one-

dimensional, because of the exceptionally strong exchange interactions along each edge. Assuming that the magnetic anisotropy (which is expected to be weak) is important only at low-temperatures, the correlation length  $\xi$  along an isolated zigzag edge is estimated[13] to be  $\sim 3000\text{\AA}/T[K]$ . Since the interedge interaction arises from an interval of  $k$ -space with width  $1/W$ , its range along the edge will be  $\sim W$ . When  $\xi$  is smaller than  $W$ , interedge interactions will have little influence. For  $\xi$  larger than  $W$ , the interedge interactions will help suppress thermal magnetization fluctuations.

*Acknowledgment*— The authors gratefully acknowledge helpful interactions with Rafi Bistritzer, Jason Hill, Hsiu-Hau Lin, Nikolai Sinitsyn, Joaquín Fernández-Rossier, Juan José Palacios and Cheng-Ching Wang. This work was supported by the Welch Foundation, NRI-SWAN, ARO, DOE and by the Spanish Ministry of Education through the MEC-Fulbright program.

\* Electronic address: jeil@physics.utexas.edu

- [1] K. Kobayashi, Phys. Rev. B, **48**, 1757 (1993).
- [2] M. Fujita et al., J. Phys. Soc. Jpn. **65**, 1920 (1996).
- [3] K. Nakada et al., Phys. Rev. B **54**, 17954 (1996).
- [4] K. Wakabayashi et al., Phys. Rev. B **59**, 8271 (1999).
- [5] K. S. Novoselov et al. Nature 438, 197-200 (2005).
- [6] M. Ezawa, Phys. Rev. B **73**, 045432 (2006).
- [7] L. Brey and H. A. Fertig, Phys. Rev. B **73**, 235411 (2006).
- [8] K.-I. Sasaki et al. J. Phys. Soc. Jpn. **75**, 074713 (2006).
- [9] T. Hikihara et al., Phys. Rev. B **68**, 035432 (2003).
- [10] Y.-W. Son et al., Phys. Rev. Lett. 97, 216803 (2006).
- [11] Y.-W. Son et al. Nature **444**, 347 (2006).
- [12] L. Pisani et al., Phys. Rev. B **75**, 064418 (2007).
- [13] O. Yazyev et al. Phys. Rev. Lett. **100**, 047209 (2008).
- [14] J. Fernández-Rossier, Phys. Rev. B **77**, 075430 (2008).
- [15] H. Lee et al., Phys. Rev. B **72**, 174431 (2005).
- [16] Malysheva and Onipko, arXiv:0802.1385v1 (2008)
- [17] K. Sasaki et al. J. Phys. Soc. Japan Vol. 77, 054703 (2008).
- [18] There are a number of reports of magnetism in damaged graphitic material which could be due to order along zigzag edge fragments. See for example H. Ohldag *et al.*, Phys. Rev. Lett. **98**, 187204 (2007).
- [19] M. Venkatesan et al., Nature 430, 630 (2004).
- [20] M. I. Katsnelson, Materials Today **10**, 20 (2007).
- [21] M. Y. Han et al, Phys. Rev. Lett. 98, 206805 (2007).
- [22] X. Li et al., Science **319**, 1229 (2008).
- [23] S. S. Datta et al. Nano Lett. **8** 1912 (2008).
- [24] J. Jung et al. in preparation.